Male sex in houseflies is determined byMdmd, a paralog of the generic splice factor geneCWC22

Sharma, Akash ; Heinze, Svenia D ; Wu, Yanli ; Kohlbrenner, Tea ; Morilla, Ian ; Brunner, Claudia ; Wimmer, Ernst A ; van de Zande, Louis ; Robinson, Mark D ; Beukeboom, Leo W ; Bopp, Daniel

Abstract: Across species, animals have diverse sex determination pathways, each consisting of a hierarchical cascade of genes and its associated regulatory mechanism. Houseflies have a distinctive polymorphic sex determination system in which a dominant male determiner, the M-factor, can reside on any of the chromosomes. We identified a gene, Musca domesticamale determiner (Mdmd), as the M-factor. Mdmd originated from a duplication of the spliceosomal factor gene CWC22 (nucampholin). Targeted Mdmd disruption results in complete sex reversal to fertile females because of a shift from male to female expression of the downstream genes transformer and doublesex The presence of Mdmd on different chromosomes indicates that Mdmd translocated to different genomic sites. Thus, an instructive signal in sex determination can arise by duplication and neofunctionalization of an essential splicing regulator.

DOI: https://doi.org/10.1126/science.aam5498

Originally published at:
Sharma, Akash; Heinze, Svenia D; Wu, Yanli; Kohlbrenner, Tea; Morilla, Ian; Brunner, Claudia; Wimmer, Ernst A; van de Zande, Louis; Robinson, Mark D; Beukeboom, Leo W; Bopp, Daniel (2017). Male sex in houseflies is determined byMdmd, a paralog of the generic splice factor geneCWC22. Science, 356(6338):642-645.
DOI: https://doi.org/10.1126/science.aam5498

Male sex in houseflies is determined by a paralog of the generic splice factor CWC22

Authors: Akash Sharma[1,2†], Svenia D. Heinze[2†], Yanli Wu[1,3], Tea Kohlbrenner[2], Ian Morilla[2,5], Claudia Brunner[2], Ernst A. Wimmer[3], Louis van de Zande[1], Mark D. Robinson[2,4], Leo W. Beukeboom[1], Daniel Bopp[2]*

† These authors contributed equally to this work

Affiliations:

[1] Groningen Institute for Evolutionary Life Sciences, University of Groningen, P.O. Box 11103, 9700 CC, Groningen, the Netherlands.

[2] Institute of Molecular Life Sciences, University of Zurich, CH-8057 Zurich, Switzerland.

[3] Department of Developmental Biology, Johann-Friedrich-Blumenbach-Institute of Zoology and Anthropology, Göttingen Center of Molecular Biosciences, University of Goettingen, Göttingen, Germany.

[4] SIB Swiss Institute of Bioinformatics, University of Zurich, Zurich, Switzerland.

[5] current address: INSERM, UMR1149, Team «Intestinal Inflammation», Research Centre of Inflammation BP 416, F-75018, Paris, France.

*Correspondence to: daniel.bopp@imls.uzh.ch

Abstract: Sex determination pathways consist of a hierarchical cascade of genes, with a bewilderingly diversity of primary signals. Houseflies have a unique polymorphic sex determination system in which a dominant male determiner *M* either resides on the Y chromosome or on different autosomes. We identified a *Musca domestica* male determiner (*Mdmd*) gene originating from a duplication of the spliceosomal factor CWC22/nucampholin, as *M*. Targeted *Mdmd* disruption results in complete sex reversal to fertile females due to a shift from male to female expression mode of downstream pathway components. Presence of *Mdmd* on the Y and on autosomes carrying *M* indicates that *Mdmd* translocated to different genomic sites. Thus an instructive signal in sex determination can arise by duplication and neofunctionalization of an essential splicing regulating gene.

One Sentence Summary: Polymorphic sex determination in houseflies results from translocation of a male-determining paralog of splice regulator CWC22.

Main Text: Genetic mechanisms for sex determination are not conserved among organismal groups. Insects illustrate this diversity, where systems vary at the chromosomal, gene and gene-regulation level between species (*1-3*). For example, under male heterogamety (XX-XY system) sex can be determined by a dominant Y-linked gene or by X-chromosome dosage (*4*). The insect sex determination pathway is conserved at the key transductory axis, the *transformer* (*tra*) and *doublesex* (*dsx*) genes, but highly diverse for the upstream instructive signals (*5-7*). The polymorphic sex determination system of the housefly, *Musca domestica*, reflects this diversity in regulation and genes (*8-11*). Males can carry a dominant male-determiner (*M*-factor) on the X or Y chromosome or any of the five autosomes (*10, 12-15*).

M acts as the instructive signal for male development in the housefly. It regulates *transformer* (*Md-tra*), a binary switch that directs female differentiation when active and male differentiation when inactive. *Md-tra* is regulated at the splicing level. The active state of *Md-tra* is initially established by maternally provided *Md-tra*. Once activated, zygotic *Md-tra* will perpetuate its female promoting function by a positive splicing feedback loop throughout development. Paternally inherited *M* prevents this maternal activation of the zygotic *Md-tra* self-regulatory loop. Early embryonic presence of male-specific splice products of *Md-tra* indicates that this regulation already starts at the cellular blastoderm stage (*11*).

We hypothesized that *M* encodes a product present only in early male embryos to prevent establishment of *Md-tra* function. Exploiting *Musca* genetics, we isolated and sequenced RNA from unisexual embryos (fig. S1). Amongst the top 14 male-specifically expressed sequences that were absent in the female *M. domestica* genome assembly (*16*), we identified five orphan contigs of the same transcription unit (Fig. 1A and table S1), which we termed *Mdmd* (for *Musca domestica male determiner*). Subsequent analysis revealed that these sequences are present in males that carry an *M*-factor on chromosome Y, II, III, or V (Fig. 1B). RT-PCR amplification confirmed exclusive presence of *Mdmd* transcripts in male embryos (Fig. 1C). Zygotic *Mdmd* transcripts first appear in 2-3h old embryos (cellularized blastoderm stage) coinciding with the first zygotic expression of *Md-tra* (*11*). *Mdmd* expression is then maintained throughout male development until adulthood (Fig. 1D). *Mdmd* encodes a protein with high homology to Complexed with Cef-1/Nucampholin (CWC22/Ncm), a spliceosome-associated protein that is required for pre-mRNA splicing and exon-junction complex (EJC) assembly (*17*). A BLAST survey of *Mdmd* over female genome scaffolds (*16*) identified a paralog (LOC101896466) of *Mdmd* structurally closely related to *ncm* genes of other insect species. In contrast to *Mdmd, ncm*

is present and expressed in both sexes (Fig. 1, B, C and D). Based on its high sequence identity to the *ncm* gene in *Drosophila* and its conserved synteny evidenced by linkage to *bicoid stability factor*, we refer to this autosomal gene as the ortholog of *ncm* and named it *Md-ncm*. Mdmd shares a high degree of identity with Md-Ncm in the MIF4G (85 %) and MA3 (79%) domains and flanking sequences but displays a substantial level of divergence in the amino-terminal and carboxy-terminal regions (Fig. 1A and fig. S2). Sequence alignments reveal that *Md-ncm* groups with prototype *ncm* genes of other insect species. However, the *Mdmd* sequences from different *M* strains form a distinct outgroup suggesting that after the duplication event *Mdmd* rapidly diverged from *Md-ncm* (Fig. 1E and fig. S3).

Multiple non-functional copies were found next to the *Mdmd gene* in the M^{III} genome (fig. S4). These copies may have arisen from local amplification to preserve *Mdmd* functionality in a non-recombining region (fig S4). Because of its long ORF *Mdmd* is particularly vulnerable to the accumulation of deleterious mutations. We identified a similar arrangement of multiple *Mdmd* copies in M^{II}, M^{V} and M^{Y} males (fig. S4). This suggests that the various *M* loci originated from a common ancestral *Mdmd* sequence which first locally multiplied and then translocated as a cluster to different sites in the genome (fig. S5).

Upon silencing of *Mdmd* by injecting dsRNA into syncytial embryos of different *M* strains, all of the surviving *M* carrying individuals developed externally as males, but 56 to 88% contained fully differentiated ovaries instead of testes with the notable exception of M^{I} males (Fig. 2, A, B and C, fig. S6). From this result, we infer that *Mdmd* is essential for specifying the male gonadal and germline fate, which is consistent with genetic findings that *M* and its target *Md-tra* govern the sexual identity of both soma and germ line (*11*). Incomplete feminization may be explained by the transient nature of embryonic RNAi. A 70% reduction of *Mdmd* transcript levels is observed in

M^{III}/+ embryos 10h after dsRNA injection, whereas after 20h levels are comparable to those in control individuals suggesting a recovery of *Mdmd* expression (Fig. 2D and E). As substantial levels of *Mdmd* transcripts were also detected in non-gonadal tissues of male adults with ovaries, restored activity of *Mdmd* at late stages apparently prevented systemic female differentiation (Fig. 2E). To conclusively test whether *Mdmd* is required for overall male differentiation, loss-of-function alleles were generated in *Mdmd* coding sequences by NHEJ mediated disruption with Cas9. Upon targeting *Mdmd* in the M^{III} strain, we recovered 59 fertile males, of which at least ten sired female progeny carrying dominant markers tightly linked to the M^{III} locus, indicating loss of its male determining function (fig. S7). These *M* containing individuals are phenotypically normal fertile females (Fig. 3A). Sequence analysis confirmed that these females carry structural aberrations in the *Mdmd* cluster (Fig. 3B). Lines M32 and M36 are most informative as the lesions specifically disrupt the ORF of *Mdmd* (Fig. 3C) and only abolish the protein coding function of this *Mdmd* copy. We conclude that the *Mdmd g*ene is indispensable for normal male development and may be the only gene in the cluster providing male function. Consistent with the role of *M* as an upstream repressor of *Md-tra*, individuals that have *Mdmd* abolished by CRISPR/Cas9 exclusively express the female splice variants of *Md-tra* and *Md-dsx* (Fig. 3D).

Based on sequence similarity, we inferred that *Mdmd* is a paralog of *Md-ncm* (*CWC22*)*, which* codes for a spliceosome-associate protein indispensable for the assembly of the Exon Junction Complex (EJC) (*17,18*). The essential functions of *CWC22* are likely to be provided by *Md-ncm* as embryonic silencing of this gene leads to early lethality in both males and females (fig. S8). However, the effect of EJC on splicing is limited to certain genes (*19*). Changes in expression levels of EJC components also affect splice site selection of alternatively spliced genes (*20*). As

tra is one of the targets on which EJCs preferentially assemble in *Drosophila* (*19*), it is conceivable that *Md-ncm* plays a crucial role in splicing regulation of *Md-tra*. Because the target of *M*, *Md-tra*, is alternatively spliced, this post-transcriptional regulatory function makes *Mdmd* an excellent candidate *M*-factor. *Mdmd* may act as a direct regulator of *Md-tra* by selectively promoting the male or preventing the female splicing mode. Alternatively, the high level of sequence similarity to its paralog opens the possibility that *Mdmd* behaves as a dominant negative, interfering with the functions of *Md-ncm* in promoting female splicing of *Md-tra*. Further study needs to elucidate the precise role of this gene in *Md-tra* splicing and can contribute to a better understanding of alternative splicing regulation.

There likely exists an enormous undiscovered source of primary signal genes for sex determination in insects. Recently, two male determiners were characterised from mosquitoes, *Nix* in *Aedes aegypti* (*21*) and *Yob* in *Anopheles gambiae* (*22*). These genes show neither sequence homology to each other nor to *Mdmd*, further pointing towards the species-specific acquisition of novel male determiners in insects. Moreover, *Mdmd* appears to be absent in the *M. domestica* strain that has an *M* factor mapped to chromosome I (fig. S4), suggesting that even intraspecific variation exists at the level of the primary signal. As insect sex determination is based on alternative splicing, its role in splicing regulation may have pre-equipped *ncm* for attaining a sex determination function. The recruitment of a *CWC22* duplicate for male function may be unique for the housefly as *ncm* paralogs have thus far not been found in other higher dipterans. Our study thus demonstrates that novel genes originating from duplication and neo-functionalization can adopt critical roles in essential developmental processes.

References and Notes:

1. J. B. Pease, M. W. Hahn, Sex chromosomes evolved from independent ancestral linkage groups in winged insects. *Mol. Biol. Evol.* **29**, 1645–1653 (2012).
2. D. Bachtrog, J. E. Mank, C. L. Peichel, M. Kirkpatrick, S. P. Otto, T. L. Ashman, M. W. Hahn, J. Kitano, I. Mayrose, R. Ming, N. Perrin, L. Ross, N. Valenzuela, J. C. Vamosi, The Tree of Sex Consortium, Sex determination: Why so many ways of doing it? *PLoS Biol.* **12**, e1001899 (2014).
3. B. Vicoso, D. Bachtrog, Numerous transitions of sex chromosomes in Diptera. *PLoS Biol.* **13**, e1002078 (2015).
4. J. W. Erickson, J. J. Quintero, Indirect effects of ploidy suggest X chromosome dose, not the X:A ratio, signals sex in *Drosophila*. *PLOS Biol.* **5**, e332 (2007).
5. L. Sanchez, Sex-determining mechanisms in insects. *Int. J. Dev. Biol.* **52**, 837–856 (2008).
6. E. C. Verhulst, L. van de Zande, L. W. Beukeboom, Insect sex determination: it all evolves around transformer. *Current Opinion in Genetics & Development* **20**, 376–383 (2010).
7. D. Bopp, G. Saccone, M. Beye, Sex determination in insects: variations on a common theme. *Sexual Development* **8**, 20–28 (2014).
8. M. G. Franco, P. G. Rubini, M. Vecchi, Sex-determinants and their distribution in various populations of *Musca domestica* L. of Western Europe. *Genet. Res., Camb.* **40**, 279–293 (1982).

9. A. Dübendorfer, M. Hediger, G. Burghardt, D. Bopp, *Musca domestica*, a window on the evolution of sex-determining mechanisms in insects. *Int. J. Dev. Biol.* **46**, 75–79 (2002).
10. R. L. Hamm, R. P. Meisel, J. G. Scott, The evolving puzzle of autosomal versus Y-linked male determination in *Musca domestica*. *Genes, Genomes, Genetics* **5**, 371–384 (2015).
11. M. Hediger, C. Henggeler, N. Meier, R. Perez, G. Saccone, D. Bopp, Molecular characterization of the key switch *F* provides a basis for understanding the rapid divergence of the sex-determining pathway in the housefly. *Genetics* **184**, 155–170 (2010).
12. T. Hiroyoshi, Sex-limited inheritance and abnormal sex ratio in strains of the housefly. *Genetics* **50**, 373–385 (1964).
13. M. G. Franco, P. G. Rubini, M. Vecchi, Sex-determinants and their distribution in various populations of *Musca domestica* L. of Western Europe. *Genet. Res., Camb.* **40**, 279–293 (1982).
14. I. Denholm, M. G. Franco, P. G. Rubini, M. Vecchi, Identification of a male determinant on the X chromosome of the housefly (*Musca domestica* L.) population in South-East England. *Genet. Res., Camb.* **42**, 311–322 (1983).
15. M. Kozielska, B. Feldmeyer, I. Pen, F. J. Weissing, L. W. Beukeboom, Are autosomal sex-determining factors of the housefly (*Musca domestica)* spreading north? *Genet. Res., Camb.* **90**, 157–165 (2008).
16. J. G. Scott, W. C. Warren, L. W. Beukeboom, D. Bopp, A. G. Clark, S. D. Giers, M. Hediger, A. K. Jones, S. Kasai, C. A. Leichter, M. Li, R. P. Meisel, P. Minx, T. D. Murphy, D. R. Nelson, W. R. Reid, F. D. Rinkevich, H. M. Robertson, T. B. Sackton, D.

B. Sattelle, F. T. Nissen, C. Tomlinson, L. van de Zande, K. K. Walden, R. K. Wilson, N. Liu, Genome of the house fly, *Musca domestica* L., a global vector of diseases with adaptations to a septic environment. *Genome Biology* **15**, 466 (2014).

17. A. L. Steckelberg, V. Boehm, A. M. Gromadzka, N. H. Gehring, CWC22 connects pre-mRNA splicing and exon junction complex assembly. *Cell Reports* **2**, 454–461 (2012).
18. A. L. Steckelberg, J. Altmueller, C. Dieterich, N. H. Gehring, CWC22-dependent pre-mRNA splicing and eIF4A3 binding enables global deposition of exon junction complexes. *Nucleic Acids Res.* **43**, 4687–4700 (2015).
19. J. Saulière, N. Haque, S. Harms, I. Barbosa, M. Blanchette, H. L. Hir, The exon junction complex differentially marks spliced junctions. *Nat. Struct. Mol. Biol.* **17**, 1269–1271 (2010).
20. Z. Wang, V. Murigneux, H. L. Hir, Transcriptome-wide modulation of splicing by the exon junction complex. *Genome Biology* **15**, 551 (2014).
21. A. B. Hall, S. Basu, X. Jiang, Y. Qi, V. A. Timoshevskiy, J. K. Biedler, M. V. Sharakhova, R. Elahi, M. A. Anderson, X. G. Chen, I. V. Sharakhov, Z. N. Adelman, Z. Tu, A male-determining factor in the mosquito *Aedes aegypti*. *Science* **348**, 1268–1270 (2015).
22. E. Krzywinska, N. J. Dennison, G. J. Lycett, J. Krzywinski, A maleness gene in the malaria mosquito *Anopheles gambiae*. *Science* **353**, 67–69 (2016).
23. R. Schmidt, M. Hediger, R. Nothiger, A. Dubendorfer, The mutation *masculinizer* (*man*) defines a sex-determining gene with maternal and zygotic functions in *Musca domestica* L. *Genetic*s **145**, 173–183 (1997a).

24. M. G. Grabherr, B. J. Haas, M. Yassour, J. Z. Levin, D. A. Thompson, I. Amit, X. Adiconis, L. Fan, R. Raychowdhury, Q. Zeng, Z. Chen, E. Mauceli, N. Hacohen, A. Gnirke, N. Rhind, F. d. Palma, B. W. Birren, C. Nusbaum, K. L. Toh, N. Friedman, A. Regev, Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol.* **29,** 644–652 (2011).

25. B. Li, C. N. Dewey, RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**, 323 (2011).

26. M. D. Robinson, D. J. McCarthy, G. K. Smyth, edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).

27. D. J. McCarthy, Y. Chen, G. K. Smyth, Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res.* **40**, 4288–4297 (2012).

28. T. D. Wu, C. K. Watanabe, GMAP: a genomic mapping and alignment program for mRNA and EST sequences. *Bioinformatics* **21**, 1859–1875 (2005).

29. Johnson, M. *et al*, NCBI BLAST: a better web interface. *Nucleic Acids Res.* **36**, W5–W9 (2008).

30. W. G. Kelly, S. Xu, M. K. Montgomery, A. Fire, Distinct requirements for somatic and germline expression of a generally expressed *Caernorhabditis elegans* gene. *Genetics* **146**, 227–238 (1997).

31. C. M. A. Coelho, B. Kolevski, C. D. Walker, I. Lavagi, T. Shaw, A. Ebert, S. J. Leevers, S. J. Marygold, A genetic screen for dominant modifiers of a small-wing phenotype in

Drosophila melanogaster identifies proteins involved in splicing and translation. *Genetics* **171**, 597–614 (2005).

32. R. C. Edgar, MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004a).
33. R. C. Edgar, MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5**, 113 (2004b).
34. F. Ronquist, J. P. Huelsenbeck, MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572–1574 (2003).
35. M. Zhong, X. Wang, J. Wen, J. Cai, C. Wu, S. M. Aly, Selection of reference genes for quantitative gene expression studies in the house fly (*Musca domestica* L.) using reverse transcription quantitative real-time PCR. *Acta Biochim Biophys Sin* **45**,1069–1073 (2013).
36. C. Ramakers, J. M. Ruijter, R. H. L. Deprez, A. F. M. Moorman, Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. *Neuroscience letters* **339**, 62–66 (2003).
37. A. Burger, H. Lindsay, A. Felker, C. Hess, C. Anders, E. Chiavacci, J. Zaugg, L. M. Weber, R. Catena, M. Jinek, M. D. Robinson, C. Mosimann, Maximizing mutagenesis with solubilized CRISPR-Cas9 ribonucleoprotein complexes. *Development* **143**, 2025–2037 (2016).

Acknowledgments: We thank E. Geuverink and A. H. Rensink for technical advice, A. Meccariello and G. Saccone for providing purified Cas9 protein and M. Hediger for helpful advice

on housefly genetics. This work was supported by the Netherlands Organisation for Scientific Research (grant no. ALW 822.02.009) and an Ubbo Emmius grant to LWB, the U4 Network sponsored by the DAAD Programme "Strategic Partnerships" and the German Federal Ministry of Education and Research to Georg-August-University Göttingen, and the Swiss National Science Foundation (grant 31003A_143883) to MDR. We thank M. Schenkel, S. Visser, G. Saccone and the Groningen Evolutionary Genetics group members for fruitful discussion. The nucleotide sequences of $Mdmd^{Y}$, $Mdmd^{II}$, $Mdmd^{III}$, and $Mdmd^{V}$ have been deposited in the DDBJ/EMBL/GenBank (accessions KY020047, KY020048, KY020049 and KY020050). Deep sequencing data are available from ArrayExpress (accession E-MTAB-5080).

Fig. 1. *Mdmd* is a male-specific paralog of the *M. domestica* CWC22 ortholog, *Md-ncm*. (A) Comparison of the two paralogs, *Mdmd* and *Md-ncm*. *Mdmd* was initially identified by extending five male-specific RNA contigs (red lines). Each exon contains a highly conserved domain, MIF4G (yellow) and MA3 (blue). Nucleotide identity is indicated in percentages. **(B)** Genomic amplifications with paralog-specific end primers show that *Mdmd* sequences are present only in males of *XY*, M^{II}, M^{III} and M^{V} strains, whereas *Md-ncm* is present in both males and females (+/+) of each strain. **(C)** RT-PCR confirming presence of *Mdmd* transcripts in 1-5h old male embryos (M^{III}/+), but not in female embryos (+/+). **(D)** Developmental expression profiles of *Mdmd* and *Md-ncm* based on RT-PCR with intron spanning primers. The upper bands in both profiles correspond to unspliced RNA and/or genomic DNA contamination, while the lower bands represent spliced transcripts **(E)** Neighbor Joining phylogenetic tree (branch label: % consensus support) of *Mdmd* and *ncm/CWC22* genes (see also MrBayes tree in fig. S3B).

Fig. 2. Embryonic silencing of *Mdmd* is transient and leads to ovarian differentiation in males. (A-C) M^{III}/+ individuals injected with dsRNA against *Mdmd*. (A) Adult abdomen with male genital structures (claspers) and inside fully differentiated ovaries (ov) (B) Dissected ovaries from the same male. (C) DAPI stained ovaries containing normal cysts composed of nurse cells (nc) and egg chambers (ec). (D) Relative levels of *Mdmd* mRNA 10h and 20h after injections with dsRNA against *Mdmd* and dsRNA against *M112* control in M^{III}/+ male embryos. (E) RT-PCR analysis of *Mdmd* transcripts and female transcripts of *Md-tra* (*Md-tra*^{F}) in normal (+/+) ovaries

(ov) and (M^{III}/+) testes (tes) and in *Mdmd* dsRNA-injected (M^{III}/+) gonadectomized bodies (gb), testes (tes) and ovaries (ov).

Fig. 3. CRISPR/Cas9 induced disruption of *Mdmd* causes complete male to female transformation. (A) F1 female of line M32 with *pw*+, *bwb*+ phenotype (left), male sibling with *pw*+, *bwb*+ phenotype (middle) and female sibling with *pw*, *bwb* phenotype (right). (B) CRISPR/Cas9 targeted sites sgF3 and sgFA in *Mdmd* (red stripes). Genomic amplifications of *Mdmd* and *Md-ncm* in F1 females of lines M6, M29, M31, M32, and M36. Upper panel F1-R4 primers amplify ORF of $Mdmd^{III}$, middle panel primers 1s/1as amplify 5' region in different *Mdmd* copies, and lower panel primers *Md-ncm*. Absence of F1-R4 amplicons in M6, M29 and M31 indicates large deletions. (C) In M32 female, a deletion of 14 bp uncovers the sgF3 target site upstream of the MIF4G domain causing a frame-shift. In M36, a deletion of 146 bp removes the same target site and extends into the MIF4G domain. (D) Expression of *Md-tra* and *Md-dsx* in sex-reverted females of lines M6, M29, M30, M31 and M32. Female splice variants are absent in control males (M^{III}/+), but present in control (+/+) and $Mdmd^{\Delta}$ females. Male splice variant of *Md-dsx*, $Md\text{-}dsx^{M}$, is only detected in control males (M^{III}/+). Expression of Cytochrome P450 (CyP) was used as an internal standard.

Supplementary Materials:

Materials and Methods

Tables S1-S2

Figures S1-S8

References (*22-37*)

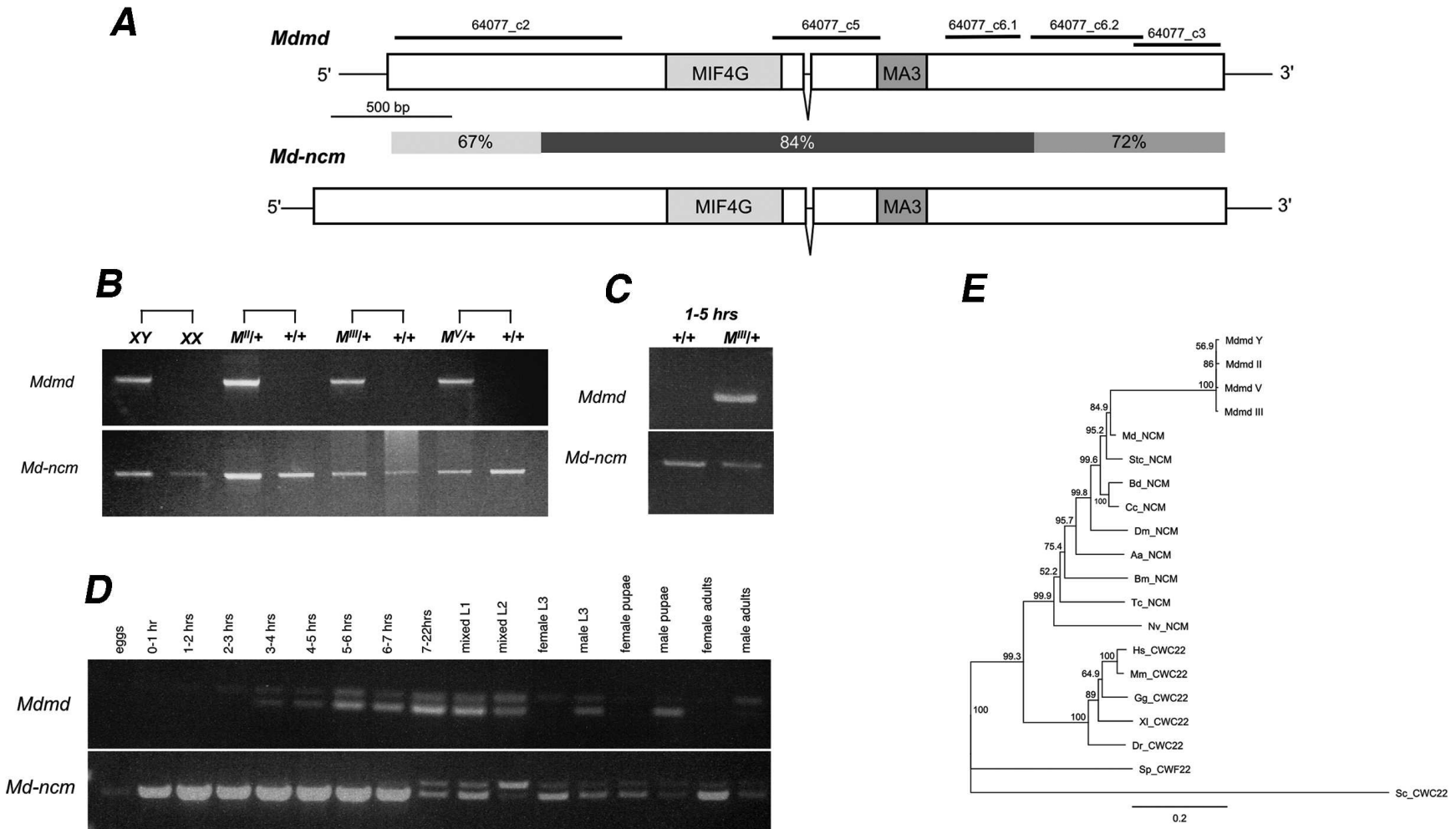
A
Mdmd
64077_c2
64077_c5
64077_c6.1
64077_c6.2
64077_c3
5'
MIF4G
MA3
3'
500 bp
67%
84%
72%
Md-ncm
5'
MIF4G
MA3
3'
B
XY
XX
MII/+
+/+
MIII/+
+/+
MV/+
+/+
Mdmd
Md-ncm
C
1-5 hrs
+/+
MIII/+
Mdmd
Md-ncm
D
eggs
0-1 hr
1-2 hrs
2-3 hrs
3-4 hrs
4-5 hrs
5-6 hrs
6-7 hrs
7-22hrs
mixed L1
mixed L2
female L3
male L3
female pupae
male pupae
female adults
male adults
Mdmd
Md-ncm
E
Mdmd Y
Mdmd II
Mdmd V
Mdmd III
Md_NCM
Stc_NCM
Bd_NCM
Cc_NCM
Dm_NCM
Aa_NCM
Bm_NCM
Tc_NCM
Nv_NCM
Hs_CWC22
Mm_CWC22
Gg_CWC22
Xl_CWC22
Dr_CWC22
Sp_CWF22
Sc_CWC22
0.2

Figure 1

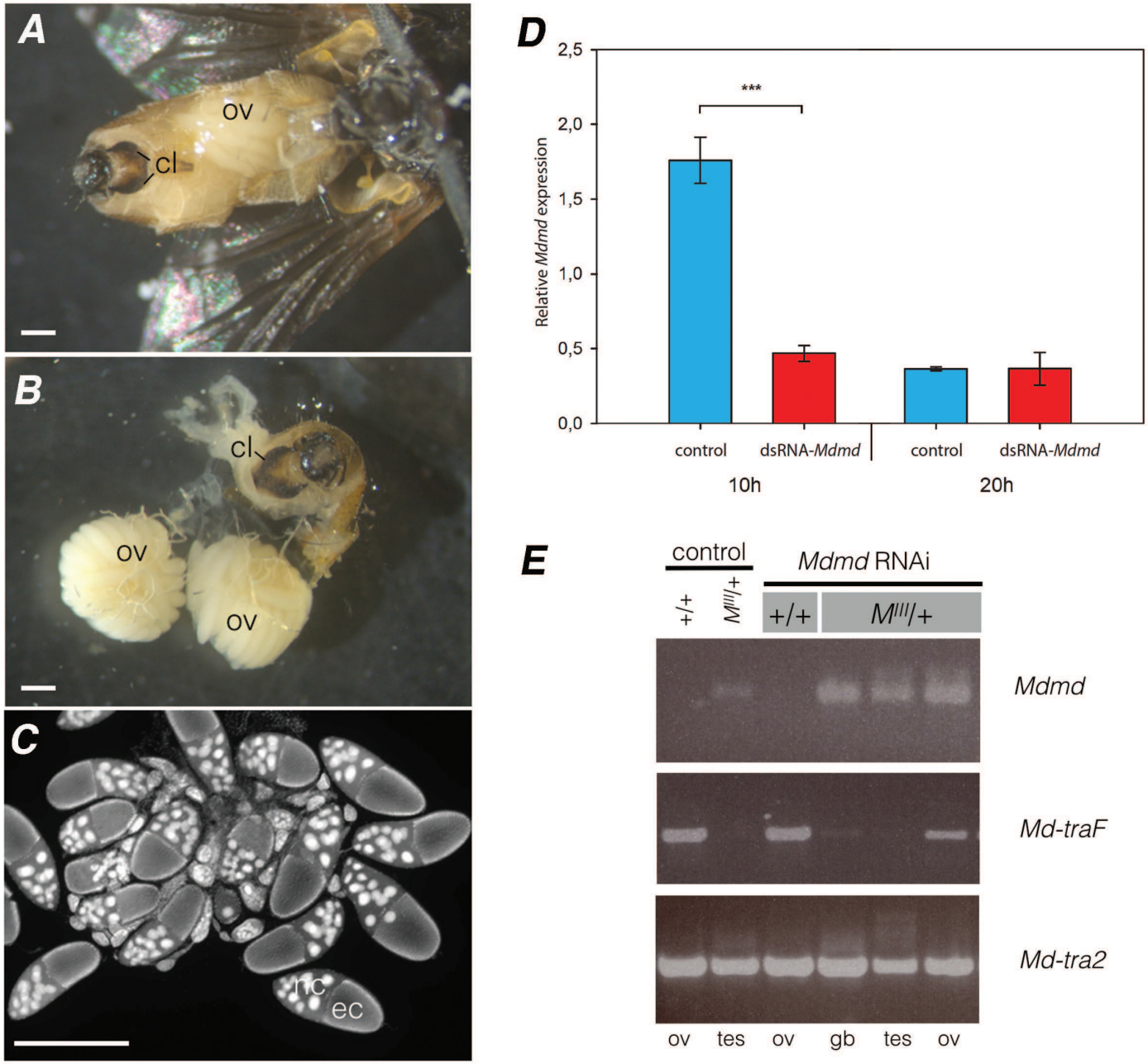
A
ov
cl
B
cl
ov
ov
C
nc
ec
D
Relative Mdmd expression
2,5
2,0
1,5
1,0
0,5
0,0
***
control
dsRNA-Mdmd
control
dsRNA-Mdmd
10h
20h
E
control
Mdmd RNAi
+/+
M^III/+
+/+
M^III/+
Mdmd
Md-traF
Md-tra2
ov
tes
ov
gb
tes
ov

Figure 2

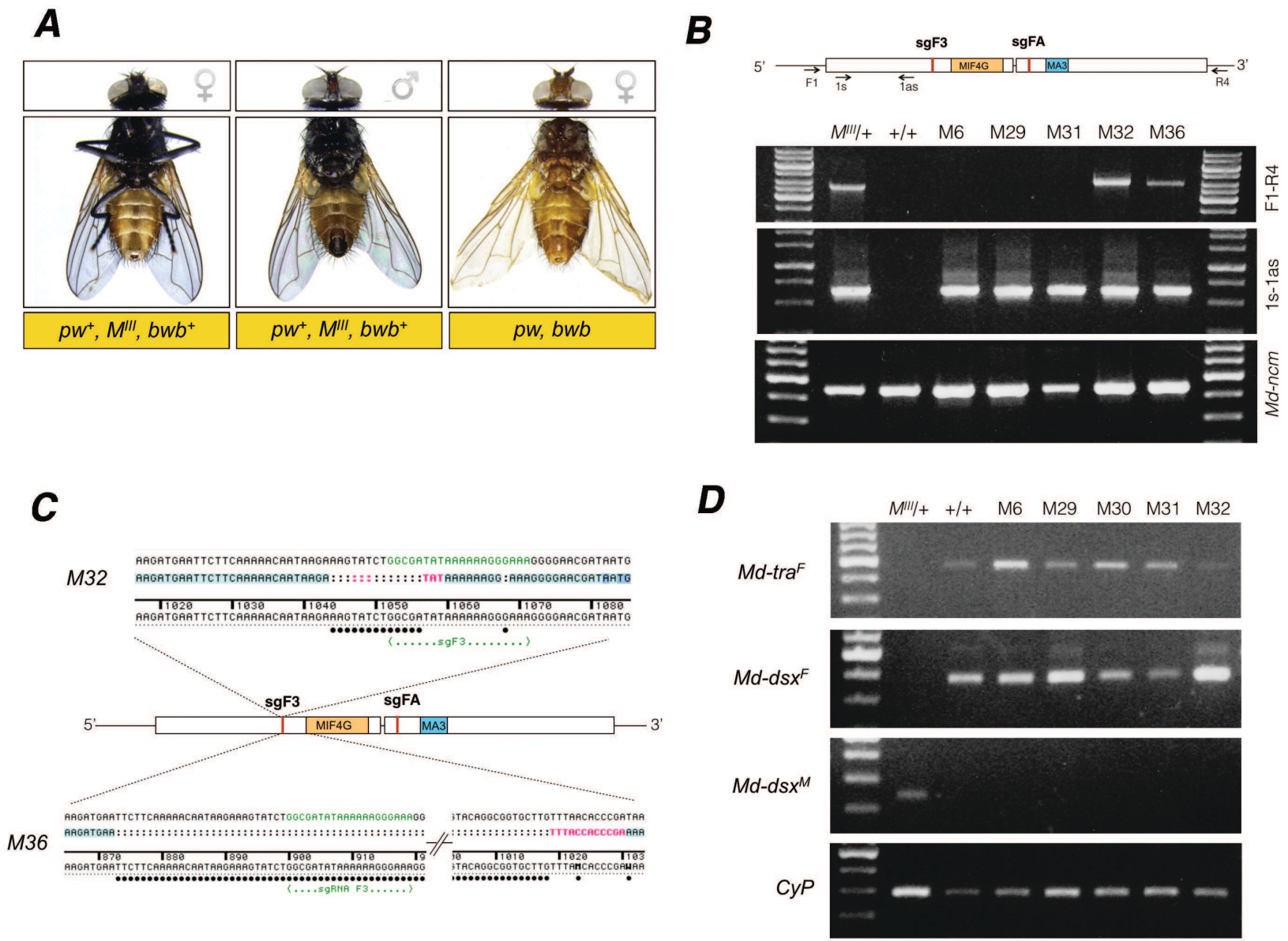
A
pw+, MIII, bwb+
pw+, MIII, bwb+
pw, bwb
B
sgF3
sgFA
5'
3'
MIF4G
MA3
F1
1s
1as
R4
MIII/+
+/+
M6
M29
M31
M32
M36
F1-R4
1s-1as
Md-ncm
C
M32
sgF3
sgFA
MIF4G
MA3
M36
D
MIII/+
+/+
M6
M29
M30
M31
M32
Md-traF
Md-dsxF
Md-dsxM
CyP

Figure 3